UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/517,615	06/10/2005	Hirotaka Suzuki	262520US6PCT	8515

22850 7590 01/03/2008
OBLON, SPIVAK, MCCLELLAND MAIER & NEUSTADT, P.C.
1940 DUKE STREET
ALEXANDRIA, VA 22314

EXAMINER	
PARK, EDWARD	
ART UNIT	**PAPER NUMBER**
2624	

NOTIFICATION DATE	DELIVERY MODE
01/03/2008	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@oblon.com
oblonpat@oblon.com
jgardner@oblon.com

Office Action Summary	Application No.	Applicant(s)
	10/517,615	SUZUKI ET AL.
	Examiner	Art Unit
	Edward Park	2624

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE _3_ MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) _1-19_ is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) _1-19_ is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on _22 December 2004_ is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☒ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _12/22/04_.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Specification

1. The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed.

Claim Objections

2. **Claim 15** is objected to because of the following informalities: It appears claim 15 has a typographical error. In regards to claim 15, in line 1, the phrase, "claim 1", appears to be a typographical error and should be corrected to "claim 11". This assumption is due to claim 10 being identical to claim 15, which depends from claim 1. For examination purposes, claim 15 will depend from claim 11.

Appropriate correction is required.

Claim Rejections - 35 USC § 102

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for

patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

4. **Claim 16** is rejected under 35 U.S.C. 102(b) as being anticipated by Schmid et al ("Local Grayvalue Invariants for Image Retrieval", IEEE).

Regarding **claim 16**, Schmid discloses an image recognition apparatus which compares an object image containing a plurality of objects with a model image containing a model to be detected and extracts the model from the object image, the apparatus comprising:

a feature point extracting step of extracting a feature point from each of the object image and the model image (see section 1.2, 2, 4.2 , interest points are local features with high information content ... database contains a set of models where each model Mk is defined by the vector of invariants Vj calculated at the interest points of the model images)

a feature quantity retention step of extracting and retaining, as a feature quantity, a density gradient direction histogram at least acquired from density gradient information in a neighboring region at the feature point in each of the object image and the model image (see figure 3, section 4.2, 4.2.1, 4.2.2, voting algorithm which is a sum of the number of times each model is selected which produces a histogram that correctly identifies the model images from the database of images);

a feature quantity comparison step of comparing each feature point of the object image with each feature point of the model image and generating a candidate-associated feature point pair having similar feature quantities (see section 4.2, 4.2.1, recognition consists of finding the model Mk

which corresponds to a given query image , that is the model which is most similar to this image .. that produces a sum that is stored in the vector T(k)); and

a model attitude estimation step of detecting the presence or absence of the model on the object image using the candidate-associated feature point pair and estimating a position and an attitude of the model (see section 4.3 geometric constraint is added based on the angle between neighbor points based on the transformation that can be locally approximated by a similarity transformation which increases the score of the object to be recognized by having it be more distinctive), if any, wherein the feature quantity comparison step itinerantly shifts one of the density gradient direction histograms of feature points to be compared in density gradient direction to find distances between the density gradient direction histograms and generates the candidate-associated feature point pair by assuming a shortest distance to be a distance between the density gradient direction histograms (see section 4.2, 4.2.1, 4.3, 4.4 semilocal constraints are utilized so there is no mis-detection of points which has the p closest features are selected which therefore transforms the vector T(k) which is determined by the distance threshold t according to the X^2 distribution).

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary

skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

6. **Claims 1-4, 8** are rejected under 35 U.S.C. 103(a) as being unpatentable over Schmid et al ("Local Grayvalue Invariants for Image Retrieval", IEEE) in view of Matsuzaki et al (US 6,804,683 B1).

Regarding **claim 1**, Schmid discloses an image recognition method which compares an object image containing a plurality of objects with a model image containing a model to be detected and extracts the model from the object image, the method comprising:

feature point extracting method for extracting a feature point from each of the object image and the model image (see section 1.2, 2, 4.2 , interest points are local features with high information content … database contains a set of models where each model Mk is defined by the vector of invariants Vj calculated at the interest points of the model images)

feature quantity retention method for extracting and retaining, as a feature quantity, a density gradient direction histogram at least acquired from density gradient information in a neighboring region at the feature point in each of the object image and the model image (see figure 3, section 4.2, 4.2.1, 4.2.2, voting algorithm which is a sum of the number of times each model is selected which produces a histogram that correctly identifies the model images from the database of images);

feature quantity comparison method for comparing each feature point of the object image with each feature point of the model image and generating a candidate-associated feature point pair having similar feature quantities (see section 4.2, 4.2.1, recognition consists of finding the model

Mk which corresponds to a given query image , that is the model which is most similar to this image .. that produces a sum that is stored in the vector T(k)); and

model attitude estimation method for detecting the presence or absence of the model on the object image using the candidate-associated feature point pair and estimating a position and an attitude of the model (see section 4.3 geometric constraint is added based on the angle between neighbor points based on the transformation that can be locally approximated by a similarity transformation which increases the score of the object to be recognized by having it be more distinctive), if any, wherein the feature quantity comparison method itinerantly shifts one of the density gradient direction histograms of feature points to be compared in density gradient direction to find distances between the density gradient direction histograms and generates the candidate-associated feature point pair by assuming a shortest distance to be a distance between the density gradient direction histograms (see section 4.2, 4.2.1, 4.3, 4.4 semilocal constraints are utilized so there is no mis-detection of points which has the p closest features are selected which therefore transforms the vector T(k) which is determined by the distance threshold t according to the X^2 distribution). While Schmid discloses these steps, Schmid does not disclose an apparatus implementing these steps.

Matsuzami, in the same field of endeavor, teaches an apparatus implementing these steps (see figure 2 numeral 2, similar image retrieving engine).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the steps of Schmid reference to utilize an apparatus as taught by Matsuzami, in order to ensure a high computational speed, and to provide the ability to isolate and extract

model images to be disseminated and used by the millions of people who have access to computers.

Regarding **claims 2-4**, Schmid further discloses extracts and retains, as the feature quantity, an average density gradient vector for each of plurality of partial regions into which the neighboring region is further divided (see section 3.1, V represent the average luminance), and the feature quantity comparison means generates the candidate-associated feature point pair based on a distance between density gradient direction histograms for the feature points to be compared and on similarity between feature vectors which are collected in the neighboring region as average density gradient vectors in each of the partial regions (see section 4.2, 4.2.1, 4.3, 4.4 semilocal constraints are utilized so there is no mis-detection of points which has the p closest features are selected which therefore transforms the vector T(k) which is determined by the distance threshold t according to the X^2 distribution);

generates a provisional candidate-associated feature point pair based on a distance between the density gradient direction histograms for the feature points to be compared and, based on the similarity between feature vectors, selects the candidate-associated feature point pair from the provisional candidate-associated feature point pair (see section 4.2, 4.3 essentially the provisional candidate implies repeating the process which is evident in any algorithm);

using a rotation angle equivalent to a shift a amount giving the shortest distance to correct a density gradient direction of a density gradient vector in the neighboring region and selects the candidate-associated feature point pair from the provisional candidate-associated feature point pair based on similarity between the feature vectors in a corrected neighboring region (see

figures 4, 5, section 4.3 geometric constraint is added based on the angel between neighbor points).

Regarding **claim 8**, Schmid further discloses candidate-associated feature point pair selection means for creating a rotation angle histogram concerning a rotation angle equivalent to a shift amount giving the shortest distance and selects a candidate-associated feature point pair giving a rotation angle for a peak in the rotation angle histogram from the candidate-associated feature point pair generated by the feature quantity comparison means (see figures 4, 5, section 4.3 geometric constraint is added based on the angel between neighbor points), wherein the model attitude estimation means detects the presence or absence of the model on the object image using a candidate-associated feature point pair selected by the candidate-associated feature point pair selection means and estimates a position and an attitude of the model, if any (see section 4.3 geometric constraint is added based on the angle between neighbor points based on the transformation that can be locally approximated by a similarity transformation which increases the score of the object to be recognized by having it be more distinctive).

7. **Claims 5-7, 9, 10** are rejected under 35 U.S.C. 103(a) as being unpatentable over Schmid et al ("Local Grayvalue Invariants for Image Retrieval", IEEE) with Matsuzaki et al (US 6,804,683 B1), and further in view of Lowe ("Object Recognition from Local Scale-Invariant Features", Computer Vision).

Regarding **claim 5-7,** Schmid with Matsuzaki combination discloses all elements as mentioned above in claim 1. Schmid with Matsuzaki combination does not disclose projecting an affine transformation parameter determined from three randomly selected candidate-associated feature point pairs onto a parameter space and finds an affine transformation

parameter to determine a position and an attitude of the model based on an affine transformation parameter belonging to a cluster having the largest number of members out of clusters formed on a parameter space, a centroid for the cluster having the largest number of members to be an affine transformation parameter to determine a position and an attitude of the model, and a least squares estimation to find an affine transformation parameter for determining a position and attitude of the model.

Lowe teaches projecting an affine transformation parameter determined from three randomly selected candidate-associated feature point pairs onto a parameter space (see section 1 scale-invariant features are efficiently identified by using a staged filter approach .. the features achieve partial invariance to local variations using affine or 3D projections by blurring the image gradient locations .. when at least 3 keys agree on the model parameters with low residual) and finds an affine transformation parameter to determine a position and an attitude of the model based on an affine transformation parameter belonging to a cluster having the largest number of members out of clusters formed on a parameter space (see sections 3, 6, 9 solve for the affine transformation parameters ... select key locations at maxima and minima of a difference of Gaussian function applied in scale space), a centroid for the cluster having the largest number of members to be an affine transformation parameter to determine a position and an attitude of the model (see section 5, cluster reliable model hypotheses is to use the Hough transform to search for keys that agree upon a particular model pose where each model key in the database contains a record of the key's parameters relative to the model coordinate system and therefore can predict the model location), and a least squares estimation to find an affine transformation parameter for determining a position and attitude of the model (see section 1, collection of keys that agree on a

potential model pose are identified and then through a least-squares fit to a final estimate of model parameters).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify Schmid with Matsuzaki combination to utilize affine transformation parameter with centroid and least squares estimation as taught by Lowe, to "allow for more accurate verification and pose determination than in approaches that rely only on indexing" (see section 9).

Regarding **claim 9 ad 10**, Schmid with Matsuzaki combination discloses all elements as mentioned above in claim 1. Schmid with Matsuzaki combination does not disclose a candidate associated feature point pair selection means for performing generalized Hough transform for a candidate-associated feature point pair generated by the feature quantity comparison means, assuming a rotation angle, enlargement and reduction ratios, and horizontal and vertical linear displacements to be a parameter space, and selecting a candidate-associated feature point pair having voted for the most voted parameter from candidate-associated feature point pairs generated by the feature quantity comparison means,

wherein the model attitude estimation means detects the presence or absence of the model on the object image using a candidate-associated feature point pair selected by the candidate-associated feature point pair selection means and estimates a position and an attitude of the model, if any; and extracting a local maximum point or a local minimum point in second-order differential filter output images with respective resolutions as the feature point, i.e., a point free from positional changes due to resolution changes within a specified range in a multi-resolution pyramid

structure acquired by repeatedly applying smoothing filtering and reduction resampling to the object image or the model image.

Lowe, in the same field of endeavor, teaches a candidate associated feature point pair selection means for performing generalized Hough transform for a candidate-associated feature point pair generated by the feature quantity comparison means, assuming a rotation angle, enlargement and reduction ratios, and horizontal and vertical linear displacements to be a parameter space, and selecting a candidate-associated feature point pair having voted for the most voted parameter from candidate-associated feature point pairs generated by the feature quantity comparison means (see section 5, 6 Hough transform to search for keys that agree upon a particular model pose … affine rotation, scale, and stretch)

wherein the model attitude estimation means detects the presence or absence of the model on the object image using a candidate-associated feature point pair selected by the candidate-associated feature point pair selection means and estimates a position and an attitude of the model, if any (see section 5-7 closest match to the correct corresponding key in the second image); and extracting a local maximum point or a local minimum point in second-order differential filter output images with respective resolutions as the feature point, i.e., a point free from positional changes due to resolution changes within a specified range in a multi-resolution pyramid structure acquired by repeatedly applying smoothing filtering and reduction resampling to the object image or the model image (see section 1 and 3, 3.1, staged filtering approach … maxima or minima of a difference of Gaussian function by building an image pyramid with resampling between each level … Gaussian kernel and its derivates are the only possible smoothing kernels for scale space analysis).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify Schmid with Matsuzaki combination to utilize a candidate-associated feature point pair and second-order differential filter as taught by Lowe, to "allow for more accurate verification and pose determination than in approaches that rely only on indexing" (see section 9).

8. **Claims 11-15** are rejected under 35 U.S.C. 103(a) as being unpatentable over Schmid et al ("Local Grayvalue Invariants for Image Retrieval", IEEE) with Lowe ("Object Recognition from Local Scale-Invariant Features", Computer Vision), and further in view of Matsuzaki et al (US 6,804,683 B1).

Regarding **claims 11-13**, Schmid discloses an image recognition method which compares an object image containing a plurality of objects with a model image containing a model to be detected and extracts the model from the object image, the apparatus comprising:

a feature point extracting step of extracting a feature point from each of the object image and the model image (see section 1.2, 2, 4.2 , interest points are local features with high information content ... database contains a set of models where each model Mk is defined by the vector of invariants Vj calculated at the interest points of the model images)

a feature quantity retention step of extracting and retaining, as a feature quantity, a density gradient direction histogram at least acquired from density gradient information in a neighboring region at the feature point in each of the object image and the model image (see figure 3, section 4.2, 4.2.1, 4.2.2, voting algorithm which is a sum of the number of times each model is selected which produces a histogram that correctly identifies the model images from the database of images);

a feature quantity comparison step of comparing each feature point of the object image with each feature point of the model image and generating a candidate-associated feature point pair having similar feature quantities (see section 4.2, 4.2.1, recognition consists of finding the model Mk which corresponds to a given query image , that is the model which is most similar to this image .. that produces a sum that is stored in the vector T(k));

a model attitude estimation step of detecting the presence or absence of the model on the object image using the candidate-associated feature point pair and estimating a position and an attitude of the model (see section 4.3 geometric constraint is added based on the angle between neighbor points based on the transformation that can be locally approximated by a similarity transformation which increases the score of the object to be recognized by having it be more distinctive).

Schmid does not teach projecting an affine transformation parameter determined from three randomly selected candidate-associated feature point pairs onto a parameter space and finds an affine transformation parameter to determine a position and an attitude of the model based on an affine transformation parameter belonging to a cluster having the largest number of members out of clusters formed on a parameter space, a centroid for the cluster having the largest number of members to be an affine transformation parameter to determine a position and an attitude of the model, and a least squares estimation to find an affine transformation parameter for determining a position and attitude of the model.

Lowe teaches projecting an affine transformation parameter determined from three randomly selected candidate-associated feature point pairs onto a parameter space (see section 1 scale-invariant features are efficiently identified by using a staged filter approach .. the features

achieve partial invariance to local variations using affine or 3D projections by blurring the image gradient locations .. when at least 3 keys agree on the model parameters with low residual) and finds an affine transformation parameter to determine a position and an attitude of the model based on an affine transformation parameter belonging to a cluster having the largest number of members out of clusters formed on a parameter space (see sections 3, 6, 9 solve for the affine transformation parameters … select key locations at maxima and minima of a difference of Gaussian function applied in scale space) and a centroid for the cluster having the largest number of members to be an affine transformation parameter to determine a position and an attitude of the model (see section 5, cluster reliable model hypotheses is to use the Hough transform to search for keys that agree upon a particular model pose where each model key in the database contains a record of the key's parameters relative to the model coordinate system and therefore can predict the model location), and a least squares estimation to find an affine transformation parameter for determining a position and attitude of the model (see section 1, collection of keys that agree on a potential model pose are identified and then through a least-squares fit to a final estimate of model parameters).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify Schmid reference to utilize affine transformation parameter with centroid and least squares estimation as taught by Lowe, to "allow for more accurate verification and pose determination than in approaches that rely only on indexing" (see section 9).

While Schmid discloses these steps, Schmid does not disclose an apparatus implementing these steps.

Matsuzami, in the same field of endeavor, teaches an apparatus implementing these steps (see figure 2 numeral 2, similar image retrieving engine).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the steps of Schmid with Lowe combination to utilize an apparatus as taught by Matsuzami, in order to ensure a high computational speed, and to provide the ability to isolate and extract model images to be disseminated and used by the millions of people who have access to computers.

Regarding **claims 14, 15**, Schmid with Matsuzaki combination discloses all elements as mentioned above in claim 11. Schmid with Matsuzaki combination does not disclose a candidate associated feature point pair selection means for performing generalized Hough transform for a candidate-associated feature point pair generated by the feature quantity comparison means, assuming a rotation angle, enlargement and reduction ratios, and horizontal and vertical linear displacements to be a parameter space, and selecting a candidate-associated feature point pair having voted for the most voted parameter from candidate-associated feature point pairs generated by the feature quantity comparison means,

wherein the model attitude estimation means detects the presence or absence of the model on the object image using a candidate-associated feature point pair selected by the candidate-associated feature point pair selection means and estimates a position and an attitude of the model, if any; and extracting a local maximum point or a local minimum point in second-order differential filter output images with respective resolutions as the feature point, i.e., a point free from positional changes due to resolution changes within a specified range in a multi-resolution pyramid

structure acquired by repeatedly applying smoothing filtering and reduction resampling to the object image or the model image.

Lowe, in the same field of endeavor, teaches a candidate associated feature point pair selection means for performing generalized Hough transform for a candidate-associated feature point pair generated by the feature quantity comparison means, assuming a rotation angle, enlargement and reduction ratios, and horizontal and vertical linear displacements to be a parameter space, and selecting a candidate-associated feature point pair having voted for the most voted parameter from candidate-associated feature point pairs generated by the feature quantity comparison means (see section 5, 6 Hough transform to search for keys that agree upon a particular model pose ... affine rotation, scale, and stretch)

wherein the model attitude estimation means detects the presence or absence of the model on the object image using a candidate-associated feature point pair selected by the candidate-associated feature point pair selection means and estimates a position and an attitude of the model, if any (see section 5-7 closest match to the correct corresponding key in the second image); and extracting a local maximum point or a local minimum point in second-order differential filter output images with respective resolutions as the feature point, i.e., a point free from positional changes due to resolution changes within a specified range in a multi-resolution pyramid structure acquired by repeatedly applying smoothing filtering and reduction resampling to the object image or the model image (see section 1 and 3, 3.1, staged filtering approach ... maxima or minima of a difference of Gaussian function by building an image pyramid with resampling between each level ... Gaussian kernel and its derivates are the only possible smoothing kernels for scale space analysis).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify Schmid with Matsuzaki combination to utilize a candidate-associated feature point pair and second-order differential filter as taught by Lowe, to "allow for more accurate verification and pose determination than in approaches that rely only on indexing" (see section 9).

9. **Claim 17** is rejected under 35 U.S.C. 103(a) as being unpatentable over Schmid et al ("Local Grayvalue Invariants for Image Retrieval", IEEE) in view of Lowe ("Object Recognition from Local Scale-Invariant Features", Computer Vision).

Regarding **claim 17,** Schmid discloses an image recognition method which compares an object image containing a plurality of objects with a model image containing a model to be detected and extracts the model from the object image, the apparatus comprising:

a feature point extracting step of extracting a feature point from each of the object image and the model image (see section 1.2, 2, 4.2 , interest points are local features with high information content … database contains a set of models where each model Mk is defined by the vector of invariants Vj calculated at the interest points of the model images)

a feature quantity retention step of extracting and retaining, as a feature quantity, a density gradient direction histogram at least acquired from density gradient information in a neighboring region at the feature point in each of the object image and the model image (see figure 3, section 4.2, 4.2.1, 4.2.2, voting algorithm which is a sum of the number of times each model is selected which produces a histogram that correctly identifies the model images from the database of images);

a feature quantity comparison step of comparing each feature point of the object image with each feature point of the model image and generating a candidate-associated feature point pair having similar feature quantities (see section 4.2, 4.2.1, recognition consists of finding the model Mk which corresponds to a given query image , that is the model which is most similar to this image .. that produces a sum that is stored in the vector T(k)); and

a model attitude estimation step of detecting the presence or absence of the model on the object image using the candidate-associated feature point pair and estimating a position and an attitude of the model (see section 4.3 geometric constraint is added based on the angle between neighbor points based on the transformation that can be locally approximated by a similarity transformation which increases the score of the object to be recognized by having it be more distinctive).

Schmid does not teach projecting an affine transformation parameter determined from three randomly selected candidate-associated feature point pairs onto a parameter space and finds an affine transformation parameter to determine a position and an attitude of the model based on an affine transformation parameter belonging to a cluster having the largest number of members out of clusters formed on a parameter space.

Lowe teaches projecting an affine transformation parameter determined from three randomly selected candidate-associated feature point pairs onto a parameter space (see section 1 scale-invariant features are efficiently identified by using a staged filter approach .. the features achieve partial invariance to local variations using affine or 3D projections by blurring the image gradient locations .. when at least 3 keys agree on the model parameters with low residual) and finds an affine transformation parameter to determine a position and an attitude of the model

based on an affine transformation parameter belonging to a cluster having the largest number of members out of clusters formed on a parameter space (see sections 3, 6, 9 solve for the affine transformation parameters ... select key locations at maxima and minima of a difference of Gaussian function applied in scale space).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify Schmid reference to utilize affine transformation parameter as taught by Lowe, to "allow for more accurate verification and pose determination than in approaches that rely only on indexing" (see section 9).

10. **Claim 18** is rejected under 35 U.S.C. 103(a) as being unpatentable over Watanabe et al (US 7,084,900 B1) in view of Schmid et al ("Local Grayvalue Invariants for Image Retrieval", IEEE).

Regarding **claim 18**, Watanabe discloses an autonomous robot apparatus (figure 1, col. 2, lines 37-60, wrist of a robot RB that is included in the robot system) capable of comparing an input image with a model image containing a model to be detected and extracting the model from the input image, the apparatus comprising:

image input means for imaging an outside environment to generate the input image (figure 1, numeral 20; col. 2, lines 37-60, image capturing device (camera or visual sensor) that captures an image of a stack of workpieces); and a processor (figure 3, numeral 1; col. 3, lines 3-10, robot operation programs that are performed by the processor).

Watanabe does not disclose a feature point extracting method for extracting a feature point from each of the input image and the model image;

feature quantity retention method for extracting and retaining, as a feature quantity, a density gradient direction histogram at least acquired from density gradient information in a neighboring region at the feature point in each of the input image and the model image;

feature quantity comparison method for comparing each feature point of the input image with each feature point of the model image and generating a candidate-associated feature point pair having similar feature quantities; and

model attitude estimation method for detecting the presence or absence of the model on the input image using the candidate-associated feature point pair and estimating a position and an attitude of the model, if any, wherein the feature quantity comparison method itinerantly shifts one of the density gradient direction histograms of feature points to be compared in density gradient direction to find distances between the density gradient direction histograms and generates the candidate-associated feature point pair by assuming a shortest distance to be a distance between the density gradient direction histograms.

Schmid, in the same field of endeavor, teaches a feature point extracting method for extracting a feature point from each of the input image and the model image (see section 1.2, 2, 4.2 , interest points are local features with high information content … database contains a set of models where each model Mk is defined by the vector of invariants Vj calculated at the interest points of the model images);

feature quantity retention method for extracting and retaining, as a feature quantity, a density gradient direction histogram at least acquired from density gradient information in a neighboring region at the feature point in each of the input image and the model image (see figure 3, section 4.2, 4.2.1, 4.2.2, voting algorithm which is a sum of the number of times each model is selected

which produces a histogram that correctly identifies the model images from the database of images);

feature quantity comparison method for comparing each feature point of the input image with each feature point of the model image and generating a candidate-associated feature point pair having similar feature quantities (see section 4.2, 4.2.1, recognition consists of finding the model Mk which corresponds to a given query image , that is the model which is most similar to this image .. that produces a sum that is stored in the vector T(k)); and

model attitude estimation method for detecting the presence or absence of the model on the input image using the candidate-associated feature point pair and estimating a position and an attitude of the model (see section 4.3 geometric constraint is added based on the angle between neighbor points based on the transformation that can be locally approximated by a similarity transformation which increases the score of the object to be recognized by having it be more distinctive), if any, wherein the feature quantity comparison method itinerantly shifts one of the density gradient direction histograms of feature points to be compared in density gradient direction to find distances between the density gradient direction histograms and generates the candidate-associated feature point pair by assuming a shortest distance to be a distance between the density gradient direction histograms (see section 4.2, 4.2.1, 4.3, 4.4 semilocal constraints are utilized so there is no mis-detection of points which has the p closest features are selected which therefore transforms the vector T(k) which is determined by the distance threshold t according to the X^2 distribution).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the Watanabe reference to utilize feature point extracting, feature quantity

retention, feature quantity comparison, model attitude estimation as taught by Schmid, in order to increase the reliability of the robot to track and retrieve targeted objects by improving the tracking ability of objects even if the image of the targeted object is "take from different viewpoints" or "only [a] part of [the] image is given" (see section 5.2.2.3, 5.2.2.4).

11. **Claim 19** is rejected under 35 U.S.C. 103(a) as being unpatentable over Watanabe et al (US 7,084,900 B1) with Schmid et al ("Local Grayvalue Invariants for Image Retrieval", IEEE), and further in view of Lowe ("Object Recognition from Local Scale-Invariant Features", Computer Vision).

Regarding **claim 19**, Watanabe discloses an autonomous robot apparatus (figure 1, col. 2, lines 37-60, wrist of a robot RB that is included in the robot system) capable of comparing an input image with a model image containing a model to be detected and extracting the model from the input image, the apparatus comprising:

image input means for imaging an outside environment to generate the input image (figure 1, numeral 20; col. 2, lines 37-60, image capturing device (camera or visual sensor) that captures an image of a stack of workpieces); and a processor (figure 3, numeral 1; col. 3, lines 3-10, robot operation programs that are performed by the processor).

Watanabe does not disclose a feature point extracting method for extracting a feature point from each of the input image and the model image;

feature quantity retention method for extracting and retaining, as a feature quantity, a density gradient direction histogram at least acquired from density gradient information in a neighboring region at the feature point in each of the input image and the model image;

feature quantity comparison method for comparing each feature point of the input image with each feature point of the model image and generating a candidate-associated feature point pair having similar feature quantities; and

model attitude estimation method for detecting the presence or absence of the model on the input image using the candidate-associated feature point pair and estimating a position and an attitude of the model, if any, wherein the model attitude estimation means repeatedly projects an affine transformation parameter determined from three randomly selected candidate-associated feature point pairs onto a parameter space and finds an affine transformation parameter to determine a position and an attitude of the model based on an affine transformation parameter belonging to a cluster having the largest number of members out of clusters formed on a parameter space.

Schmid, in the same field of endeavor, teaches a feature point extracting method for extracting a feature point from each of the input image and the model image (see section 1.2, 2, 4.2 , interest points are local features with high information content … database contains a set of models where each model Mk is defined by the vector of invariants Vj calculated at the interest points of the model images);

feature quantity retention method for extracting and retaining, as a feature quantity, a density gradient direction histogram at least acquired from density gradient information in a neighboring region at the feature point in each of the input image and the model image (see figure 3, section 4.2, 4.2.1, 4.2.2, voting algorithm which is a sum of the number of times each model is selected which produces a histogram that correctly identifies the model images from the database of images);

feature quantity comparison method for comparing each feature point of the input image with each feature point of the model image and generating a candidate-associated feature point pair having similar feature quantities (see section 4.2, 4.2.1, recognition consists of finding the model Mk which corresponds to a given query image , that is the model which is most similar to this image .. that produces a sum that is stored in the vector T(k)); and

model attitude estimation method for detecting the presence or absence of the model on the input image using the candidate-associated feature point pair and estimating a position and an attitude of the model (see section 4.3 geometric constraint is added based on the angle between neighbor points based on the transformation that can be locally approximated by a similarity transformation which increases the score of the object to be recognized by having it be more distinctive).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the Watanabe reference to utilize feature point extracting, feature quantity retention, feature quantity comparison, model attitude estimation as taught by Schmid, in order to increase the reliability of the robot to track and retrieve targeted objects by improving the tracking ability of objects even if the image of the targeted object is "take from different viewpoints" or "only [a] part of [the] image is given" (see section 5.2.2.3, 5.2.2.4).

Lowe teaches projecting an affine transformation parameter determined from three randomly selected candidate-associated feature point pairs onto a parameter space (see section 1 scale-invariant features are efficiently identified by using a staged filter approach .. the features achieve partial invariance to local variations using affine or 3D projections by blurring the image gradient locations .. when at least 3 keys agree on the model parameters with low residual) and

finds an affine transformation parameter to determine a position and an attitude of the model based on an affine transformation parameter belonging to a cluster having the largest number of members out of clusters formed on a parameter space (see sections 3, 6, 9 solve for the affine transformation parameters ... select key locations at maxima and minima of a difference of Gaussian function applied in scale space).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify Watanabe with Schmid combination to utilize affine transformation parameter as taught by Lowe, to "allow for more accurate verification and pose determination than in approaches that rely only on indexing" (see section 9).

Conclusion

12. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Edward Park whose telephone number is (571) 270-1576. The examiner can normally be reached on M-F 10:30 - 20:00, (EST).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Vikkram Bali can be reached on (571) 272-7415. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Edward Park
Examiner
Art Unit 2624

/Edward Park/

VIKKRAM BALI
PRIMARY EXAMINER